

DOI:10.12361/2661-3263-05-08-115564

大数据背景下统计数据资源整合分析

宋双双

对外经济贸易大学统计学院, 中国·北京 100105

【摘要】大数据的出现, 不仅为社会发展创造了大量机遇, 而且还提出了诸多挑战, 这一背景下, 想要推动现代社会向着更加良好的方向发展, 必须要注重统计数据资源的整合。基于此, 本文通过对统计数据资源整合的简单介绍, 进而分析了大数据背景下统计数据资源整合存在的问题, 之后以此为基础, 提出了相应的优化策略, 以此更好地对统计数据资源整合提供支持。

【关键词】大数据; 统计数据; 资源整合; 优化策略

Integration Analysis of Statistical Data Resources under the Background of Big Data

Song Shuangshuang

School of Statistics, University of International Business and Economics, Beijing 100105

[Abstract] The emergence of big data not only creates a lot of opportunities for social development, but also puts forward many challenges. Under this background, in order to promote the development of modern society towards a better direction, we must pay attention to the integration of statistical data resources. Based on this, this paper briefly introduces the integration of statistical data resources, and then analyzes the problems existing in the integration of statistical data resources under the background of big data. On this basis, it puts forward the corresponding optimization strategy, so as to provide better support for the integration of statistical data resources.

[Keywords] Big data; Statistical data; Resource integration; Optimization strategy

引言

信息化技术迅猛发展的今天, 社会各界逐渐出现了越来越多的统计数据, 这些统计数据当中不仅包含一些有用的信息, 而且还有很多无价值的信息, 若盲目地对统计数据进行处理, 不仅会增加相关部门与人员的工作量, 而且还会影响工作质量, 不利于现代社会的发展。所以, 在大数据背景下, 相关部门应注重统计数据资源的整合, 以此为统计数据的进一步筛选与处理提供支持, 使统计数据在现代社会发展中发挥出更大的作用。

1 统计数据资源整合概述

1.1 基本定义

统计部门工作过程中, 统计数据资源整合是一项较为重要的工作, 指的是以构建一体化、更全面、更高效的数据资源管理体系为目标, 以统计数据的元数据化为中心, 突破现有体系、工作、系统等方面存在的限制, 全面对统计数据进行分析与处理, 从而得到价值密度更高的数据集群, 以此为统计数据更好地应用打下坚实基础。其中, 元数据是重要的概念, 可对其他类型的数据进行描述, 以将各类数据的存储位置、特性、类型等准确展示出来, 有利于统计数据的查找与使用^[1]。元数据化的处理, 指的是对一套制度表与其他制度相关指标、分类等进行整理, 形成相对应的元数据集群, 之后以此为基础, 开发出相应的管理系统, 以提升元数据的管理力度。

1.2 大数据背景下统计数据资源整合的特点

大数据背景下, 统计数据资源整合工作呈现出全新的特点, 具体来说, 主要体现在下述几个方面: (1) 全员参与。当前, 统计数据资源数量庞大, 其中涉及到经济、人口、财务、管理等多方面的内容, 若仅由一方面人员开展该项工作, 则会导致工作当中出现诸多问题。针对这一情况, 则需要各方

面人员共同参与进来, 利用不同领域人员的专业特点, 全面对统计数据资源整合处理, 以保证数据整合结果的质量。(2) 信息更加丰富。不论是信息的处理方法, 还是信息输出的途径, 均呈现出多样性特点。其中, 在信息处理方法方面, 主要包括数据采集、筛选、挖掘、开发等; 在数据输出途径方面, 包括图像、表格及动态画面等。(3) 综合系统。统计数据资源整合是一项复杂的工作, 只有对其进行综合性的分析与处理, 才会真正挖掘出统计数据中有价值的信息, 为各方面人员的使用提供支持。

2 大数据背景下统计数据资源整合存在的问题

2.1 数据存储不统一

数据资源整合过程中, 数据的存储至关重要, 采取有效的方式对数据进行存储, 可以为后续数据的分析、挖掘等工作的开展提供支持。但通过大量实践证明, 现代统计数据资源存储依然存在不统一的问题, 各类型、各领域的数据分别存储在不同的位置, 缺乏专业的数据存储库, 不仅能够影响数据的整合与使用, 而且还会带来更高的安全隐患, 当计算机出现问题, 或是工作人员出现变动后, 可能导致部分数据失真, 或是管理软件出现变化, 使部分数据无法正常使用; 对于部分统计数据存储软件来说, 性能与功能较差, 不符合大数据背景下统计数据存储的要求, 也会在一定程度上影响数据的整合^[2]。

2.2 应用软件不统一

统计工作涉及到现代社会发展的方方面面, 如人口普查、经济发展水平调查等, 不同方面的统计数据之间虽然具有独立性的特点, 但各类型数据间也存在一定的联系, 只有将这些统计数据资源进行统一、全面的处理, 才会真正挖掘出数据中心包含的有价值信息, 为统计数据的应用提供支持。但对于现有处理软件来说, 通常利用不同软件对不同方面的数据进行处理,

导致得到的数据格式存在一定差异,使得数据整合效果差,只能流于表面,而难以将所有数据深度融合。此外,为了保证各方面数据能够融合到一起,需要采用多种不同的数据处理软件,大大增加了统计数据资源整合的工作量。

2.3 数据共享效果差

统计数据资源整合工作的开展,需要由统计部门各个方面人员共同参与进来,只有这样,才会提升数据整合的质量,有利于数据更好地使用。但实际当中,统计部门并未对此产生高度重视,各专业、各方面人员独自开展该项工作,未能有效对数据进行共享,导致统计部门内部出现“信息孤岛”的问题,所有数据难以融合到一起,不利于统计数据的整合与使用,从而对现代社会发展造成一定干扰。

2.4 数据开发不足

数据开发是统计数据资源整合中的重要组成部分,用于挖掘出数据中包含的有价值信息,以此为其他相关工作的开展提供支持。现代统计数据开发过程中,在处理方法、性能问题等方面均存在一定缺陷,未能采用先进的数据分析软件,大部分软件并无数据特征分析功能,难以全面对数据信息进行挖掘,无法为其他工作的开展提供良好的信息支持。

2.5 缺乏完善的数据资源整合体系

大数据背景下,统计数据资源整合工作的开展离不开整合体系的支持,只有构建出完善的数据资源整合体系,才可提升数据资源的整合效果,为统计数据的使用提供支持。然而在实际当中,统计部门依然未能制定出完善的数据整合体系,无法对该项工作的开展提供指导,导致实际经常出现各种问题。一方面,相关人员的综合素质并不是很高,大部分只是掌握统计方面的理论知识,而大数据、信息化方面的理论知识掌握得则并不是很多,在一定程度上影响统计数据的整合。另一方面,统计数据资源整合需要诸多技术提供支持,如大数据技术、信息化技术等^[3]。但在实际工作开展时,受到技术水平因素的影响,使得这些技术无法发挥出最大的作用,从而影响数据整合效果。

3 大数据背景下统计数据资源整合的优化

3.1 注重统计数据的存储管理

大数据背景下,为了进一步提升统计数据资源整合工作的开展质量,必须要注重统计数据的存储管理。针对这一情况,统计部门应构建出良好的数据存储系统,各项统计工作开展时,将所获得的数据信息录入到该存储系统内,由该系统对所有统计数据进行存储,防止统计数据出现丢失的问题。这样统一存储管理后,可加强对统计数据的保护,提升整个统计数据的安全性,以确保统计数据在实际应用中发挥出更大的作用。与此同时,注重对数据存储系统的维护,不断赋予其更多良好性能,以符合大数据背景下,统计数据存储与整合的要求。

3.2 设计统一的应用软件

由上述分析可知,现代统计部门工作过程中,未能应用统计的应用软件,在一定程度上影响了统计数据的整合处理。所以,大数据背景下,统计部门应设计出统一的应用软件,以此为统计数据的整合打下坚实基础。应用软件的设计与开发,需要聘请专业机构完成,利用软件研发机构掌握的专业技术与手段,结合统计数据资源整合的需求,开发出性能更加良好,且功能更加健全的应用软件,通过该应用软件的应用,一方面,可以得到相同格式的统计数据,促进各方面统计数据的融合,利于数据的互通,为数据的挖掘与使用提供帮助。另一方面,还可减少其他处理软件的应用,降低统计数据资源整合工作量,提升数据整合效率,以确保统计数据更好应用。

3.3 加强统计数据的共享

想要提升统计数据资源整合质量,应加强统计数据的共享。首先,各专业、岗位人员日常工作当中,应转变以往的思维理念,认识到数据共享的重要性,并以此为基础,主动与其他方面人员进行交流与沟通,及时将本岗位工作传输给其他岗位人员,并第一时间获取其他岗位产生的统计数据,以提升统计数据的完善性,为统计数据的融合做好准备。另一方面,可针对统计数据共享的需求,构建出良好的共享平台,即统一的数据管理系统,各专业、岗位产生统计数据后,均将这些数据录入到该管理系统内,各方面人员想要获取其他方面的数据后,只需要登录到该系统内,通过关键的词汇进行检索。

3.4 强化统计数据的开发

有效挖掘出数据当中包含的有价值信息,可以为数据的使用提供重要帮助。为此,在统计数据资源整合过程中,需要强化统计数据的开发,以挖掘出更多有价值的信息。想要达到这一要求,首先,优化现有数据开发方法,逐渐融入可视化技术,打造出性能更加良好的可视化开发工具,符合数据显示的多维需求,同时支持多种数据的集成方式,它不仅限于数据库,还支持数据仓库和团队协作中的数据、文本等。其次,在现有数据开发系统当中,增加数据特征挖掘的功能,以深入挖掘数据中心包含的有价值信息,以提升统计数据的应用价值。

3.5 构建完善的数据资源整合体系

除上述几个方面之外,还应构建出完善的数据资源整合体系。首先,打造出一支高素质的数据统计团队。统计部门可加强与周边各大高校的合作,利用高校专业的人才培养手段,定向为统计部门培养统计工作人员,确保这些人员掌握统计相关理论知识的同时,还应对信息化技术具有准确的了解,使得他们在统计数据资源整合时,有效将信息化技术融入进来,以提升数据整合与使用的质量。其次,加强对科学技术的研发,开发出更多与数据资源整合相关的技术手段,并将这些技术手段应用到统计数据资源整合工作当中,利用这些更加先进的科学技术,以加强对统计数据资源的整合,为更好地对统计数据应用打下坚实基础。

4 总结

综上所述,统计是现代社会发展中的重要工作之一,有效开展该项工作,可得到诸多与现代社会发展相关的国情调查数据,通过对这些统计数据的整合与分析,可以掌握国情发展具体情况,以此为未来发展决策的制定提供指导。但需要注意的是,受到各方面因素的影响,导致现代统计数据资源整合工作中依然存在较多问题,在一定程度上影响现代社会的发展。为此,统计部门应针对实际问题,不断对统计数据资源整合工作予以优化,以此为社会发展提供更大的帮助。

参考文献:

- [1]陈晶,管林杰,张钟海等.长江大保护数据资源整合与数据中台建设研究——以湖北省为例[J].人民长江,2022,53(2):198-202.
- [2]张龙,崔应寿,马骁等.基于GIS的交通运输数据资源整合共享与查询系统架构研究[J].交通世界,2021,04(23):25-289.
- [3]张伟伟,乐恒,田冉等.自然资源与规划数据资源整合研究——以宁波为例[J].测绘与空间地理信息,2020,43(11):112-115.

作者简介:

宋双双,(1987-)女,汉,山东菏泽人,大学本科,研究方向:大数据科学与技术。对外经济贸易大学统计学院在职人员高级课程研修班学员。